Subspace methods

When to use

- Local minima
- Reasonable initial guess
- High dimensionality

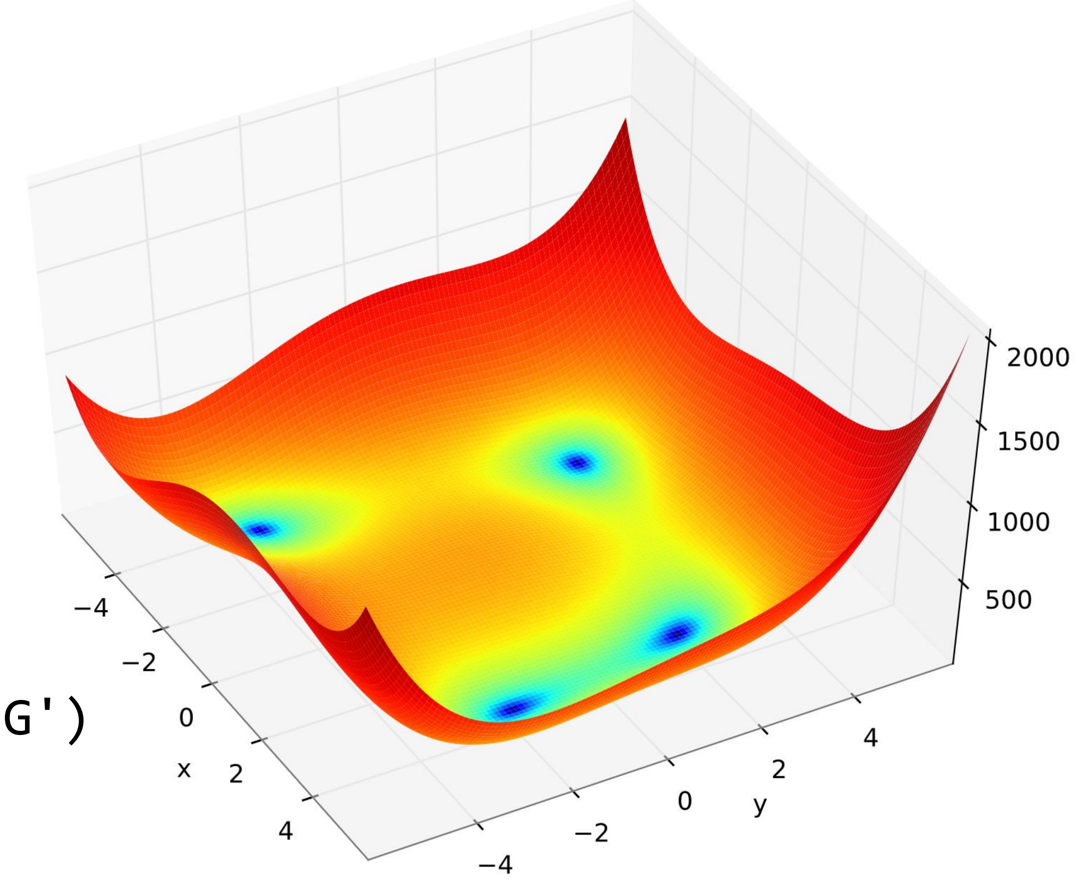
When not to use

- Noisy function evaluations

Popular representatives

- Conjugate Gradients

```
        scipy.optimize.minimize(method='CG')
```



CG (Fletcher-Reeves)

Initialise

$$p_0 = -\nabla f(a_n)$$

Line search

$$\alpha_{n+1} = \arg\min f(a_n + \alpha p_n)$$

Update optimisation

$$a_{n+1} = a_n + \alpha_{n+1} p_n$$

New problem-orthogonal search direction

$$\beta_{n+1} = \frac{||\nabla f(a_{n+1})||^2}{||\nabla f(a_n)||^2}$$

$$p_{n+1} = -\nabla f(a_{n+1}) + \beta_{n+1} p_n$$

CG

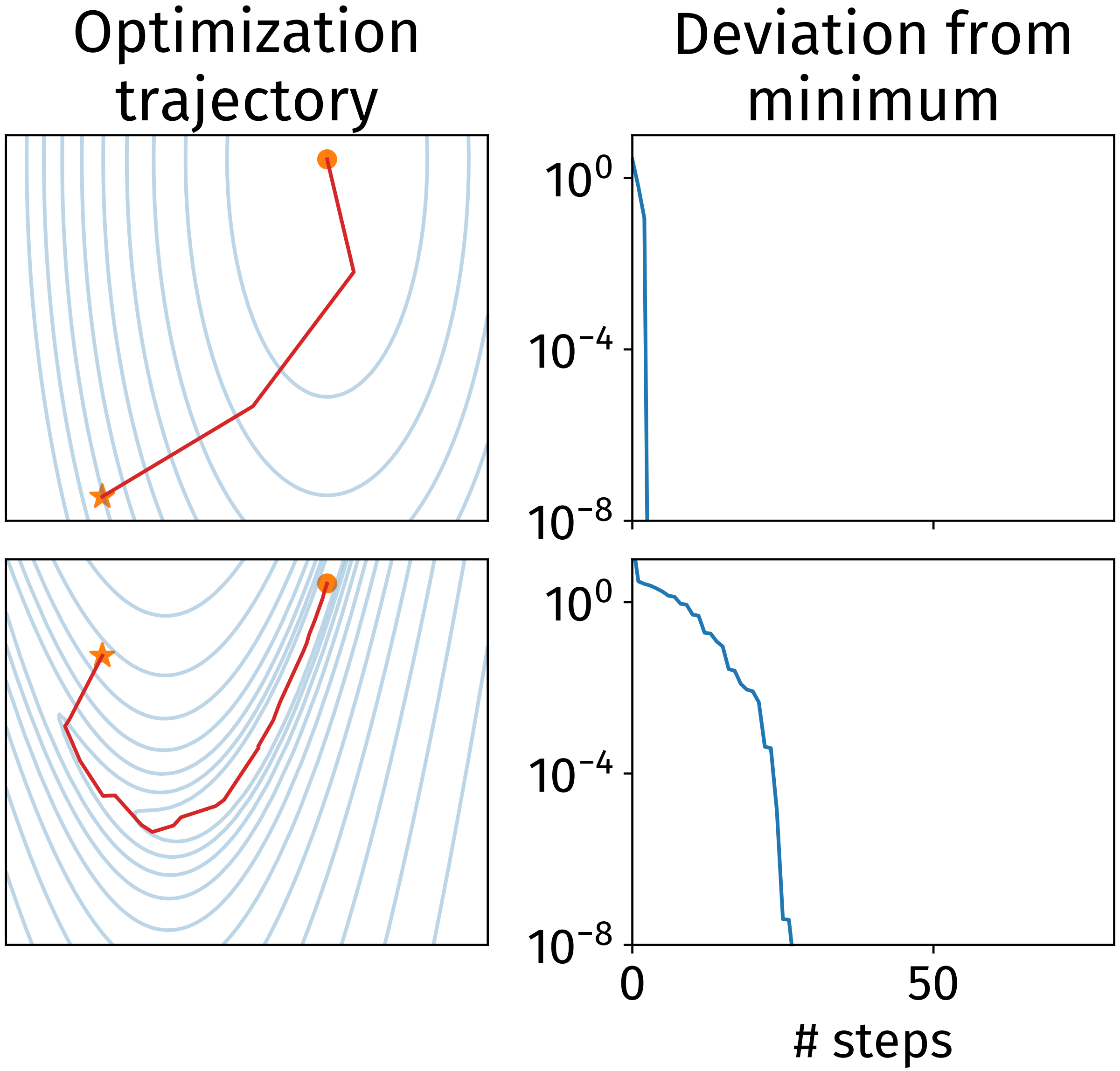
Subsequent residual minimisation

Variants

- Other search directions β

Problems

- Numerical stability: restart



Stochastic optimization family

When to use

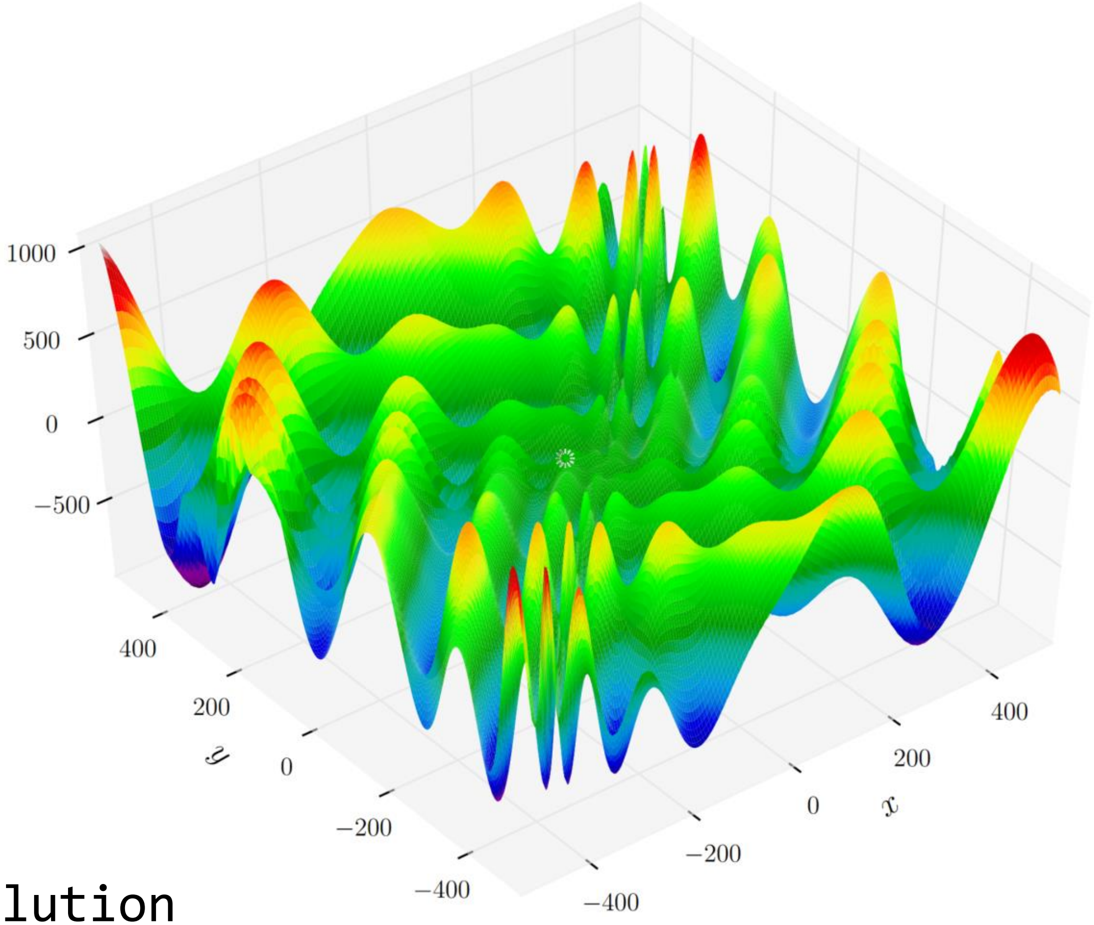
- Large domain
- Highly non-linear
- Small attractive basins
- Many minima
- High dimensionality

When not to use

- (Cheap) gradients available

Popular representatives

- Simulated annealing
  `scipy.optimize.basinhopping`
- Genetic algorithms
  `scipy.optimize.differential_evolution`



Optimization: Caveats

Convergence

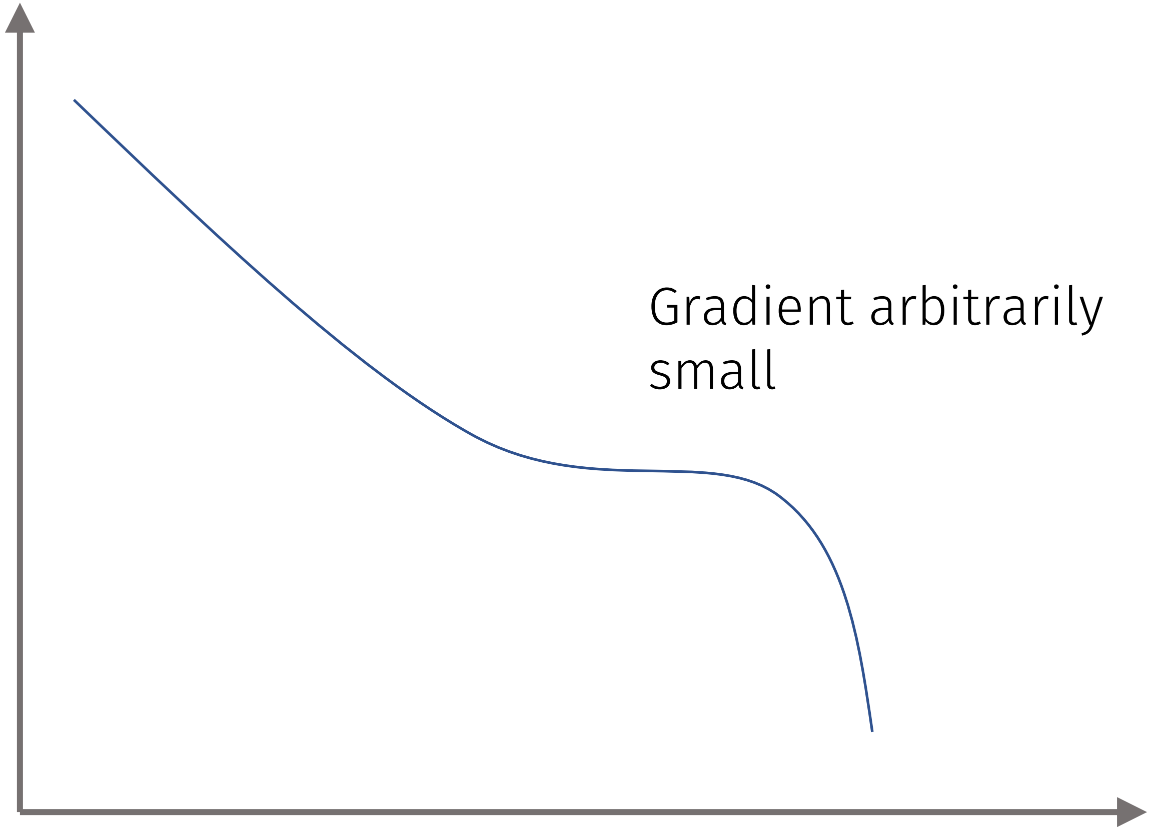
- Hard to establish
- Gradient necessary, but not sufficient
- Hessian expensive
- Local property

Numerical stability

- Finite differences
- Conjugate Gradients
- Shallow minima

Cost of Hessians

- Scales as N^2
    - Water: N=9
    - Caffeine: N=72
- Often only from finite differences



Gradient arbitrarily small

Optimization: Caveats

Curse of dimensionality

- Search space quickly increases
- Often forces tiny optimization steps

Preconditioning

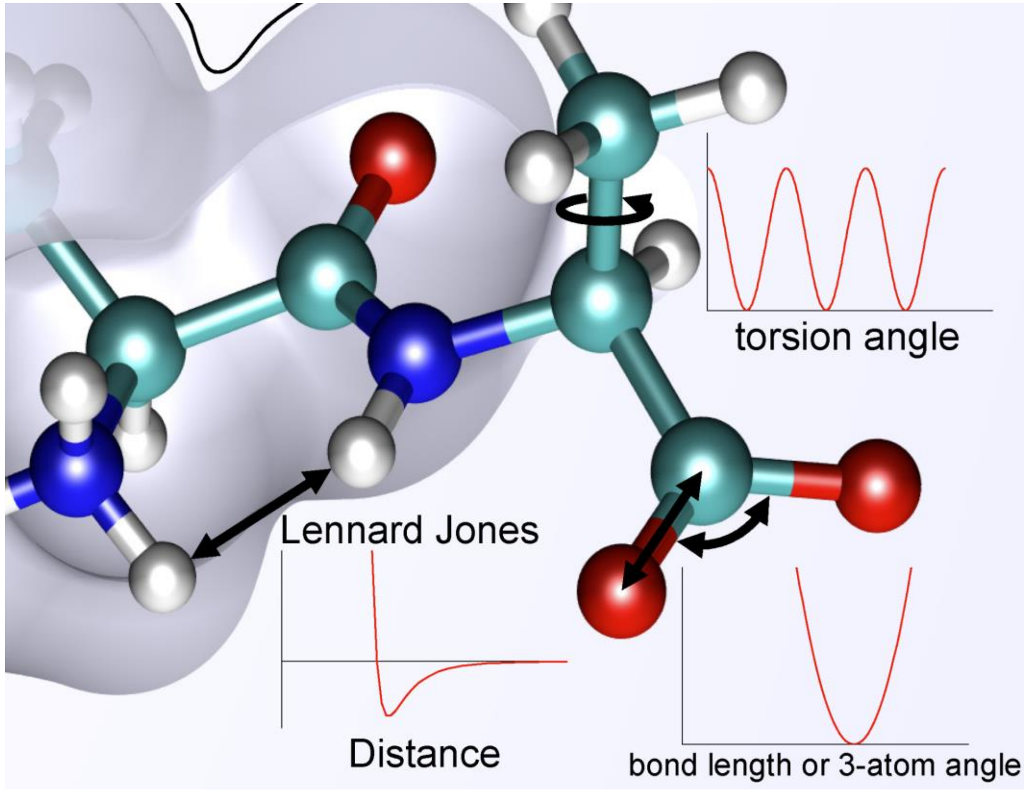
- Math not equal to finite-precision implementations
- Transform problem into an equivalent one
- Focus on numerical stability
- Key: use libraries when possible or implement algorithms verbatim

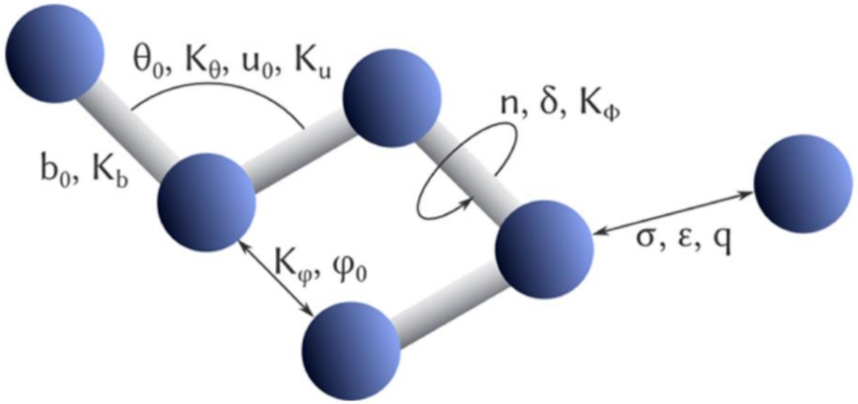
Families of Approximations

Classical molecular dynamics

- Typically fixed bonds
- No quantum effects
- Reference: quantum data

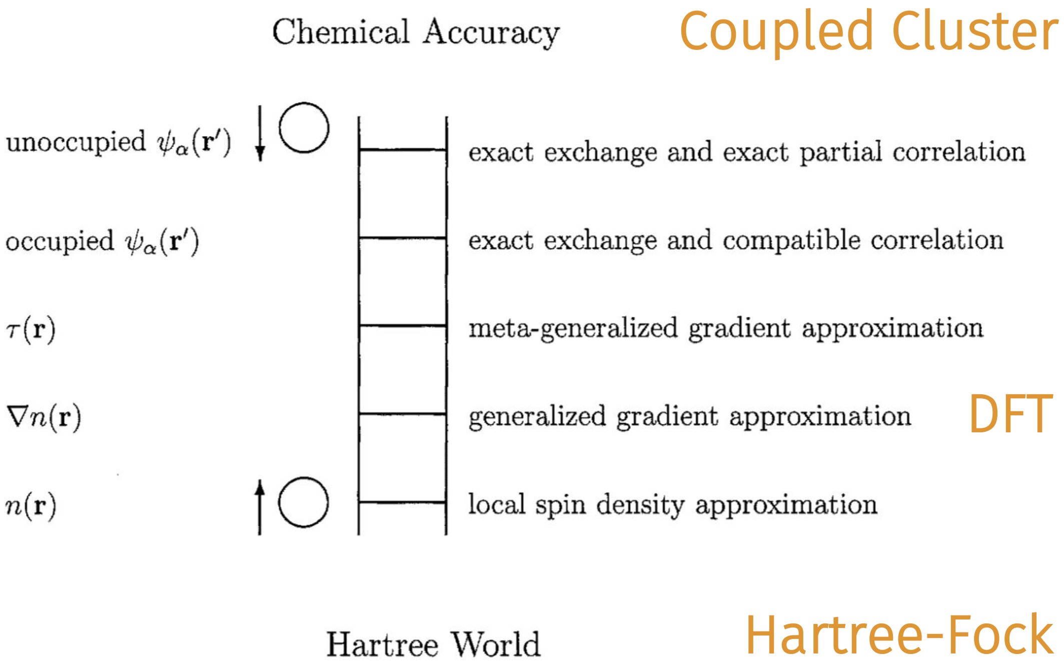
$$
\begin{aligned}
E = &\sum_{\text{bonds}} K_b(b - b_0)^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_0)^2 \\
&+ \sum_{\text{dihedrals}} K_\phi(1 + \cos(n\phi - \delta)) \\
&+ \sum_{\text{improper}} K_\varphi(\varphi - \varphi_0)^2 + \sum_{\text{Urey-Bradley}} K_u(u - u_0)^2 \\
&+ \sum_{i<j} 4\varepsilon\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_{i<j} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}
\end{aligned}
$$




θ_0, K_θ, u_0, K_u
n, δ, K_Φ
b_0, K_b
K_φ, φ_0
σ, ε, q

https://dx.doi.org/10.1021/jp507464m / wikicommons

Quantum chemistry



Chemical Accuracy

Coupled Cluster

unoccupied $\psi_\alpha(\mathbf{r}')$ ↓ ○ exact exchange and exact partial correlation

occupied $\psi_\alpha(\mathbf{r}')$ exact exchange and compatible correlation

$\tau(\mathbf{r})$ meta-generalized gradient approximation

$\nabla n(\mathbf{r})$ generalized gradient approximation

DFT

$n(\mathbf{r})$ ↑ ○ local spin density approximation

Hartree World

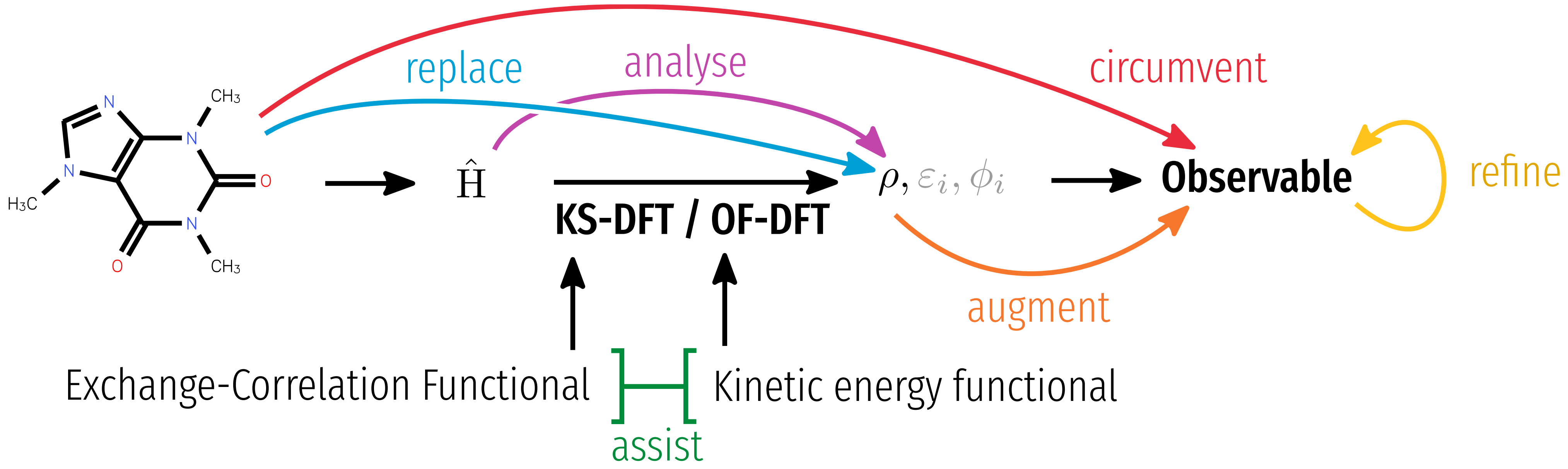
Hartree-Fock

FIGURE 1. Jacob's ladder of density functional approximations. Any resemblance to the Tower of Babel is purely coincidental. Also shown are angels in the spherical approximation, ascending and descending. Users are free to choose the rungs appropriate to their accuracy requirements and computational resources. However, at present their safety can be guaranteed only on the two lowest rungs.

John Perdew

Michael Willmann, 1691

Machine learning force fields



replace

analyse

circumvent

$\hat{\mathrm{H}}$

KS-DFT / OF-DFT

$\rho, \varepsilon_i, \phi_i$

Observable

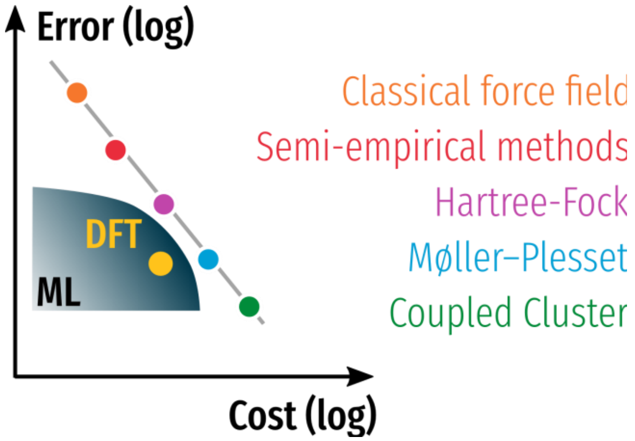
refine

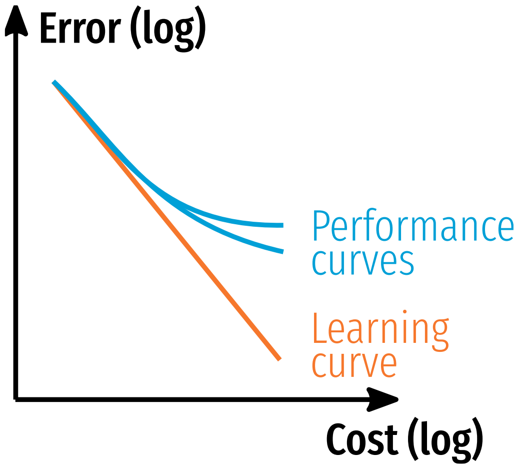
augment

Exchange-Correlation Functional

Kinetic energy functional

assist


Error (log)
Cost (log)
DFT
ML
Classical force field
Semi-empirical methods
Hartree-Fock
Møller–Plesset
Coupled Cluster


Error (log)
Cost (log)
Performance curves
Learning curve